



0590
0320

RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/940,925
Source: OIPE
Date Processed by STIC: 3-15-02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER
VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND
TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name,
Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
Or
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two,
2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office,
Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Raw Sequence Listing Error Summary .

ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/940,925

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 ☒ **Wrapped Nucleics
Wrapped Aminos** The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

- 2 ☐ **Invalid Line Length** The rules require that a line not exceed 72 characters in length. This includes white spaces.

- 3 ☐ **Misaligned Amino
Numbering** The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

- 4 ☐ **Non-ASCII** The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

- 5 ☐ **Variable Length** Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

- 6 ☐ **PatentIn 2.0
"bug"** A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

- 7 ☐ **Skipped Sequences
(OLD RULES)** Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
 (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
 (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
 This sequence is intentionally skipped

 Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

- 8 ☐ **Skipped Sequences
(NEW RULES)** Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
 <210> sequence id number
 <400> sequence id number
 000

- 9 ☐ **Use of n's or Xaa's
(NEW RULES)** Use of n's and/or Xaa's have been detected in the Sequence Listing.
 Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
 In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

- 10 ☐ **Invalid <213>
Response** Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

- 11 ☐ **Use of <220>** Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
 Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
 (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)

- 12 ☐ **PatentIn 2.0
"bug"** Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.



**Does Not Comply
Corrected Diskette Needed**

OIIPE

RAW SEQUENCE LISTING DATE: 03/15/2002
PATENT APPLICATION: US/09/940,925 TIME: 14:41:44

Input Set : A:\Seq-sub.app
Output Set: N:\CRF3\03152002\I940925.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:
6 (i) APPLICANT: BROW, MARY ANN D.
7 LYAMICHEV, VICTOR I.
8 OLIVE, DAVID M.
10 (ii) TITLE OF INVENTION: RAPID DETECTION AND IDENTIFICATION OF
11 PATHOGENS
13 (iii) NUMBER OF SEQUENCES: 165
15 (iv) CORRESPONDENCE ADDRESS:
16 (A) ADDRESSEE: MEDLEN & CARROLL
17 (B) STREET: 220 MONTGOMERY STREET, SUITE 2200
18 (C) CITY: SAN FRANCISCO
19 (D) STATE: CALIFORNIA
20 (E) COUNTRY: UNITED STATES OF AMERICA
21 (F) ZIP: 94104
23 (v) COMPUTER READABLE FORM:
24 (A) MEDIUM TYPE: Floppy disk
25 (B) COMPUTER: IBM PC compatible
26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
29 (vi) CURRENT APPLICATION DATA:
C--> 30 (A) APPLICATION NUMBER: US/09/940,925
C--> 31 (B) FILING DATE: 28-Aug-2001
32 (C) CLASSIFICATION:
34 (viii) ATTORNEY/AGENT INFORMATION:
35 (A) NAME: CARROLL, PETER G.
36 (B) REGISTRATION NUMBER: 32,837
37 (C) REFERENCE/DOCKET NUMBER: FORS-01756
39 (ix) TELECOMMUNICATION INFORMATION:
40 (A) TELEPHONE: (415) 705-8410.
41 (B) TELEFAX: (415) 397-8338

ERRORED SEQUENCES

44 (2) INFORMATION FOR SEQ ID NO: 1:
46 (i) SEQUENCE CHARACTERISTICS:
47 (A) LENGTH: 2506 base pairs
48 (B) TYPE: nucleic acid
49 (C) STRANDEDNESS: double
50 (D) TOPOLOGY: linear
52 (ii) MOLECULE TYPE: DNA (genomic)
56 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/940,925

DATE: 03/15/2002
TIME: 14:41:44

Input Set : A:\Seq-sub.app
Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 58 ATGAGGGGGA TGCTGCCCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT
59 GGACGGCCAC 60
E--> 61 CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG
62 GGGGAGCCG 120
E--> 64 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTCAAGGA
65 GGACGGGGAC 180
E--> 67 GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC
68 CTACGGGGGG 240
E--> 70 TACAAGGCGG GCCGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC
71 CCTCATCAAG 300
E--> 73 GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA
74 GGCGGACGAC 360
E--> 76 GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG
77 CATCCTCACC 420
E--> 79 GCCGACAAAG ACCTTTACCA GTCCTTTCC GACCGCATCC ACGTCCTCCA
80 CCCCAGAGGG 480
E--> 82 TACCTCATCA CCCCAGGCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCGA
83 CCAAGTGGGCC 540
E--> 85 GACTACCGGG CCCTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA
86 GGGCATCGGG 600
E--> 88 GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGAAGCCCT
89 CCTCAAGAAC 660
E--> 91 CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA
92 CGATCTGAAG 720
E--> 94 CTCTCCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA
95 CTTGCGCAA 780
E--> 97 AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA
98 GTTGGCAGC 840
E--> 100 CTCTCCACG AGTTCCGGCT TCTGGAAAGC CCCAAGGCC TGGAGGAGGC
101 CCCCTGGCCC 900
E--> 103 CCGCCGGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCCGCA AGGAGCCCAT
104 GTGGGCCGAT 960
E--> 106 CTTCTGGCCC TGGCCGCCG CAGGGGGGGC CGGGTCCACC GGGCCCCCGA
107 GCCTTATAAA 1020
E--> 109 GCCCTCAGGG ACCTGAAGGA GGC GCGGGGG CTTCTCGCCA AAGACCTGAG
110 CGTCTGGCC 1080
E--> 112 CTGAGGGAAG GCCTTGGCCT CCCGCCCGG GACGACCCCA TGCTCCTCGC
113 CTACCTCCTG 1140
E--> 115 GACCCTTCCA ACACCACCCC CGAGGGGGTG GCCCGGCGCT ACGGCGGGGA
116 GTGGACGGAG 1200
E--> 118 GAGGCGGGGG AGCGGGCCGC CCTTTCCGAG AGGCTCTTCG CCAACCTGTG
119 GGGGAGGCTT 1260
E--> 121 GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCTT
122 TTCCGCTGTC 1320
E--> 124 CTGGCCACA TGGAGGCCAC GGGGGTGCGC CTGGACGTGG CCTATCTCAG
125 GCCTTGTC 1380
E--> 127 CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCCGAGG TCTTCCGCCT
128 GGCCGGCCAC 1440
E--> 130 CCCTTCAACC TCAACTCCCC GGACCAGCTG GAAAGGTCC TCTTTGACGA

```

← Wrapped Text,
see item # 1 on
ERROR Summary
Sheet.

↓
The type of errors shown exist throughout
the Sequence Listing. Please check subsequent
sequences for similar errors.

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

131 GCTAGGGCTT      1500
E--> 133 CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGCTCCA CCAGCGCCGC
134 CGTCCTGGAG      1560
E--> 136 GCCCTCCGCG AGGCCCACCC CATCGTGGAG AAGATCCTGC AGTACCGGGA
137 GCTACCAAG       1620
E--> 139 CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCCAGGAC
140 GGGCCGCCTC      1680
E--> 142 CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGGCAGGC TAAGTAGCTC
143 CGATCCCAAC      1740
E--> 145 CTCCAGAACA TCCCCGTCGG CACCCCGCTT GGGCAGAGGA TCCGCCGGGC
146 CTTTCATCGCC     1800
E--> 148 GAGGAGGGGT GGCTATTGGT GGCCCTGGAC TATAGCCAGA TAGAGCTCAG
149 GGTGCTGGCC      1860
E--> 151 CACCTCTCCG GCGACGAGAA CCTGATCCGG GTCTTCCAGG AGGGGCGGGA
152 CATCCACACG      1920
E--> 154 GAGACCGCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCT
155 GATGCGCCGG      1980
E--> 157 GCGGCCAAGA CCATCAACTT CGGGGTCCCTC TACGGCATGT CGGCCACCG
158 CCTCTCCCAG      2040
E--> 160 GAGCTAGCCA TCCCTTACGA GGAGGCCCAG GCCTTCATTG AGCGCTACTT
161 TCAGAGCTTC      2100
E--> 163 CCAAGGTGC GGGCCTGGAT TGAGAAGACC CTGGAGGAGG GCAGGAGGCG
164 GGGGTACGTG      2160
E--> 166 GAGACCCTCT TCGGCCGCCG CCGCTACGTG CCAGACCTAG AGGCCCGGGT
167 GAAGAGCGTG      2220
E--> 169 CGGGAGGCGG CCGAGCGCAT GGCCTTCAAC ATGCCCCGTC AGGGCACCGC
170 CGCCGACCTC      2280
E--> 172 ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC
173 CAGGATGCTC      2340
E--> 175 CTTCAGGTCC ACGACGAGCT GGTCTCTGAG GCCCCAAAAG AGAGGGCGGA
176 GGCCGTGGCC      2400
E--> 178 CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCCTGG CCGTGCCCT
179 GGAGGTGGAG      2460
E--> 181 GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC
W--> 182 2506
184 (2) INFORMATION FOR SEQ ID NO: 2:
186 (i) SEQUENCE CHARACTERISTICS:
187 (A) LENGTH: 2496 base pairs
188 (B) TYPE: nucleic acid
189 (C) STRANDEDNESS: double
190 (D) TOPOLOGY: linear
192 (ii) MOLECULE TYPE: DNA (genomic)
196 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
E--> 198 ATGGCGATGC TTCCCTCTT TGAGCCCAA GGCCGCGTGC TCCTGGTGGA
199 CGGCCACCAC      60
E--> 201 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCCTACCA CCAGCCGCGG
202 CGAACCCGTT      120
E--> 204 CAGGCGGTCT ACGGCTTCGC CAAAAGCCTC CTCAAGGCC TGAAGGAGGA
205 CGGGGACGTG      180

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

E--> 207 GTGGTGGTGG TCTTTGACGC CAAGGCCCCC TCCTTCCGCC ACGAGGCCTA
208 CGAGGCCTAC 240
E--> 210 AAGGCGGGCC GGGCCCCAC CCCGGAGGAC TTTCCCCGGC AGCTGGCCCT
211 CATCAAGGAG 300
E--> 213 TTGGTGGACC TCCTAGGCCT TGTGCGGCTG GAGGTTCCTG GCTTTGAGGC
214 GGACGACGTG 360
E--> 216 CTGGCCACCC TGGCCAAGCG GCGGAAAAG GAGGGGTACG AGGTGCGCAT
217 CCTCACTGCC 420
E--> 219 GACCGCGACC TCTACCAGCT CCTTTCGGAG CGCATCGCCA TCCTCCACCC
220 TGAGGGGTAC 480
E--> 222 CTGATCACCC CGGCGTGGCT TTACGAGAAG TACGGCCTGC GCGCGGAGCA
223 GTGGGTGGAC 540
E--> 225 TACCGGGCCC TGGCGGGGA CCCCTCGGAT AACATCCCCG GGGTGAAGGG
226 CATCGGGGAG 600
E--> 228 AAGACCGCCC AGAGGCTCAT CCGCGAGTGG GGGAGCCTGG AAAACCTCTT
229 CCAGCACCTG 660
E--> 231 GACCAGGTGA AGCCCTCCTT GCGGGAGAAG CTCCAGGCGG GCATGGAGGC
232 CCTGGCCCTT 720
E--> 234 TCCCGGAAGC TTTCCCAAGT GCACACTGAC CTGCCCCCTG AGGTGGACTT
235 CGGGAGGCGC 780
E--> 237 CGCACACCCA ACCTGGAGGG TCTGCGGGCT TTTTGGAGC GGTGGAGTT
238 TGGAAGCCTC 840
E--> 240 CTCCACGAGT TCGGCCTCCT GGAGGGGCGG AAGGCGGCAG AGGAGGCCCC
241 CTGGCCCCCT 900
E--> 243 CCGGAAGGGG CTTTTTTGGG CTTTTCTTTT TCCCGTCCCG AGCCCATGTG
244 GGCCGAGCTT 960
E--> 246 CTGGCCCTGG CTGGGGCGTG GGAGGGGCGC CTCCATCGGG CACAAGACCC
247 CCTTAGGGGC 1020
E--> 249 CTGAGGGACC TTAAGGGGGT GCGGGGAATC CTGGCCAAGG ACCTGGCGGT
250 TTTGGCCCTG 1080
E--> 252 CGGGAGGGCC TGGACCTCTT CCCAGAGGAC GACCCCATGC TCCTGGCCTA
253 CCTTCTGGAC 1140
E--> 255 CCCTCCAACA CCACCCCTGA GGGGGTGGCC CGGCGTTACG GGGGGGAGTG
256 GACGGAGGAT 1200
E--> 258 GCGGGGGAGA GGGCCCTCCT GGCCGAGCGC CTCTTCCAGA CCCTAAAGGA
259 GCGCCTTAAG 1260
E--> 261 GGAGAAGAAC GCCTGCTTTG GCTTTACGAG GAGGTGGAGA AGCCGCTTTC
262 CCGGGTGTG 1320
E--> 264 GCGCGGATGG AGGCCACGGG GGTCCGGCTG GACGTGGCCT ACCTCCAGGC
265 CCTCTCCCTG 1380
E--> 267 GAGGTGGAGG CGGAGGTGCG CCAGCTGGAG GAGGAGGTCT TCCGCTGGC
268 CGCCACCCC 1440
E--> 270 TTCAACCTCA ACTCCGCGA CCAGCTGGAG CGGGTGCTCT TTGACGAGCT
271 GGGCCTGCCT 1500
E--> 273 GCCATCGGCA AGACGGAGAA GACGGGGAAA CGCTCCACCA GCGCTGCCGT
274 GCTGGAGGCC 1560
E--> 276 CTGCGAGAGG CCCACCCAT CGTGGACCGC ATCCTGCAGT ACCGGGAGCT
277 CACCAAGCTC 1620
E--> 279 AAGAACACCT ACATAGACCC CCTGCCCCGCC CTGGTCCACC CCAAGACCGG

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

      280 CCGGCTCCAC      1680
E--> 282 ACCCGCTTCA ACCAGACGGC CACCGCCACG GGCAGGCTTT CCAGCTCCGA
      283 CCCCACCTG      1740
E--> 285 CAGAACATCC CCGTGCGCAC CCCTCTGGGC CAGCGCATCC GCCGAGCCTT
      286 CGTGGCCGAG      1800
E--> 288 GAGGGCTGGG TGCTGGTGGT CTTGGACTAC AGCCAGATTG AGCTTCGGGT
      289 CCTGGCCAC      1860
E--> 291 CTCTCCGGGG ACGAGAACCT GATCCGGGTC TTTCAGGAGG GGAGGGACAT
      292 CCACACCCAG      1920
E--> 294 ACCGCCAGCT GGATGTTTCGG CGTTTCCCCC GAAGGGGTAG ACCCTCTGAT
      295 GCGCCGGGCG      1980
E--> 297 GCCAAGACCA TCAACTTCGG GGTGCTCTAC GGCATGTCCG CCCACCGCCT
      298 CTCCGGGGAG      2040
E--> 300 CTTTCCATCC CCTACGAGGA GCGGGTGGCC TTCATTGAGC GCTACTTCCA
      301 GAGCTACCCC      2100
E--> 303 AAGGTGCGGG CCTGGATTGA GGGGACCCTC GAGGAGGGCC GCCGGCGGGG
      304 GTATGTGGAG      2160
E--> 306 ACCCTCTTCG GCCGCCGGCG CTATGTGCCC GACCTCAACG CCCGGGTGAA
      307 GAGCGTGC GC      2220
E--> 309 GAGGCGGCGG AGCGCATGGC CTTCAACATG CCGGTCCAGG GCACCGCCGC
      310 CGACCTCATG      2280
E--> 312 AAGCTGGCCA TGGTGCGGCT TTTCCCCCGG CTTCAGGAAC TGGGGGCGAG
      313 GATGCTTTTG      2340
E--> 315 CAGGTGCACG ACGAGCTGGT CCTCGAGGCC CCCAAGGACC GGGCGGAGAG
      316 GGTAGCCGCT      2400
E--> 318 TTGGCCAAGG AGGTCATGGA GGGGGTCTGG CCCCTGCAGG TGCCCCCTGGA
      319 GGTGGAGGTG      2460
      321 GGCCTGGGGG AGGACTGGCT CTCCGCCAAG GAGTAG
      323 (2) INFORMATION FOR SEQ ID NO: 3:
      325 (i) SEQUENCE CHARACTERISTICS:
      326 (A) LENGTH: 2504 base pairs
      327 (B) TYPE: nucleic acid
      328 (C) STRANDEDNESS: double
      329 (D) TOPOLOGY: linear
      331 (ii) MOLECULE TYPE: DNA (genomic)
      335 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
E--> 337 ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCGGG TCCTCCTGGT
      338 GGACGGCCAC      60
E--> 340 CACCTGGCCT ACCGCACCTT CTTGCGCCTG AAGGGCCTCA CCACGAGCCG
      341 GGGCGAACCG      120
E--> 343 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA
      344 GGACGGGTAC      180
E--> 346 AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
      347 GGCCTACGAG      240
E--> 349 GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT
      350 CGCCCTCATC      300
E--> 352 AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA
      353 CGAGGCGGAC      360
E--> 355 GACGTTCTCG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT

```

Same

2496

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

356 GCGCATCCTC      420
E--> 358 ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT
359 CCACCCCGAG      480
E--> 361 GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC
362 GGAGCAGTGG      540
E--> 364 GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT
365 CAAGGGCATC      600
E--> 367 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGAAAA
368 CCTCCTCAAG      660
E--> 370 AACCTGGACC GGGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA
371 CCTGGAAGAC      720
E--> 373 CTCAGGCTCT CCTTGGAGCT CTCCCGGGTG CGCACCAGACC TCCCCCTGGA
374 GTTGGACCTC      780
E--> 376 GCCCAGGGGC GGGAGCCCGA CCGGGAGGGG CTTAGGGCCT TCCTGGAGAG
377 GCTGGAGTTC      840
E--> 379 GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA
380 GGAGGCCCCC      900
E--> 382 TGGCCCCCGC CGGAAGGGGC CTTCGTGGGC TTCGTCTCTT CCCGCCCCGA
383 GCCCATGTGG      960
E--> 385 GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC
386 AGCAGACCCC     1020
E--> 388 TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA
389 CCTCGCCGTC     1080
E--> 391 TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCCATGCT
392 CCTCGCCTAC     1140
E--> 394 CTCCTGGACC CCTCCAACAC CACCCCGAG GGGGTGGCGC GCGCTACGG
395 GGGGGAGTGG     1200
E--> 397 ACGGAGGACG CCGCCCACCG GGCCCTCCTC TCGGAGAGGC TCCATCGGAA
398 CCTCCTTAAG     1260
E--> 400 CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCACG AGGTGGAAAA
401 GCCCCTCTCC     1320
E--> 403 CGGGTCCTGG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA
404 CCTTCAGGCC     1380
E--> 406 CTTTCCCTGG AGCTTGCGGA GGAGATCCGC CGCCTCGAGG AGGAGGTCTT
407 CCGCTTGGCG     1440
E--> 409 GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGAAA GGGTGCTCTT
410 TGACGAGCTT     1500
E--> 412 AGGCTTCCCG CCTTGGGGAA GACGCAAAAAG ACAGGCAAGC GCTCCACCAG
413 CGCCGCGGTG     1560
E--> 415 CTGGAGGCCC TACGGGAGGC CCACCCATC GTGGAGAAGA TCCTCCAGCA
416 CCGGGAGCTC     1620
E--> 418 ACCAAGCTCA AGAACACCTA CGTGGACCCC CTCCCAAGCC TCGTCCACCC
419 GAGGACGGGC     1680
E--> 421 CGCCTCCACA CCCGCTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG
422 TAGCTCCGAC     1740
E--> 424 CCAACCTGCG AGAACATCCC CGTCCGCACC CCCTTGGGCC AGAGGATCCG
425 CCGGGCCTTC     1800
E--> 427 GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA
428 GCTCCGCGTC     1860

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 430 CTCGCCCACC TCTCCGGGGA CGAAAACCTG ATCAGGGTCT TCCAGGAGGG
      431 GAAGGACATC      1920
E--> 433 CACACCCAGA CCGCAAGCTG GATGTTCTGGC GTCCCCCCGG AGGCCGTGGA
      434 CCCCTGATG      1980
E--> 436 CGCCGGGCGG CCAAGACGGT GAACTTCGGC GTCCTCTACG GCATGTCCGC
      437 CCATAGGCTC      2040
E--> 439 TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC
      440 TACTTCCAAA      2100
E--> 442 GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCCTGGA GGAGGGGAGG
      443 AAGCGGGGCT      2160
E--> 445 ACGTGGAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC
      446 CGGGTGAAGA      2220
E--> 448 GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC
      449 ACCGCCGCCG      2280
E--> 451 ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCCGCCT CCGGGAGATG
      452 GGGGCCCCGA      2340
E--> 454 TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG
      455 GCCGAGGAGG      2400
E--> 457 TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCTATCC CCTCGCCGTG
      458 CCCCTGGAGG      2460
E--> 460 TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCCAAGGG TTAG
W--> 461 2504
      977 (2) INFORMATION FOR SEQ ID NO: 7:
      979 (i) SEQUENCE CHARACTERISTICS:
      980 (A) LENGTH: 2502 base pairs
      981 (B) TYPE: nucleic acid
      982 (C) STRANDEDNESS: single
      983 (D) TOPOLOGY: linear
      985 (ii) MOLECULE TYPE: DNA (genomic)
      989 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
E--> 991 ATGNNGGCGA TGCTTCCCT CTTTGAGCCC AAAGGCCGGG TCCTCCTGGT
      992 GGACGGCCAC      60
E--> 994 CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA CCACCAGCCG
      995 GGGCGAACCG      120
E--> 997 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA
      998 GGACGGGGAC      180
E--> 1000 NNGGCGGTGN TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
      1001 GGCCTACGAG      240
E--> 1003 GCCTACAAGG CGGGCCGGGC CCCCACCCCG GAGGACTTTC CCCGGCAGCT
      1004 CGCCCTCATC      300
E--> 1006 AAGGAGCTGG TGGACCTCCT GGGGCTTGCG CGCCTCGAGG TCCCCGGCTA
      1007 CGAGGCGGAC      360
E--> 1009 GACGTNCTGG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT
      1010 GCGCATCCTC      420
E--> 1012 ACCGCCGACC GCGACCTCTA CCAGCTCCTT TCCGACCGCA TCGCCGTCTT
      1013 CCACCCCGAG      480
E--> 1015 GGGTACCTCA TCACCCCGGC GTGGCTTTGG GAGAAGTACG GCCTGAGGCC
      1016 GGAGCAGTGG      540
E--> 1018 GTGGACTACC GGGCCCTGGC GGGGGACCCC TCCGACAACC TCCCCGGGGT

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1019 CAAGGGCATC      600
E--> 1021 GGGGAGAAGA CCGCCCNAG GCTCCTCNAG GAGTGGGGGA GCCTGGAAAA
1022 CCTCCTCAAG      660
E--> 1024 AACCTGGACC GGGTGAAGCC CGCCNTCCGG GAGAAGATCC AGGCCCACAT
1025 GGANGACCTG      720
E--> 1027 ANGCTCTCCT GGGAGCTNTC CCAGGTGCGC ACCGACCTGC CCCTGGAGGT
1028 GGACTTCGCC      780
E--> 1030 AAGNGGCGGG AGCCCGACCG GGAGGGGCTT AGGGCCTTTC TGGAGAGGCT
1031 GGAGTTTGGC      840
E--> 1033 AGCCTCCTCC ACGAGTTCGG CCTCCTGGAG GGCCCCAAGG CCCTGGAGGA
1034 GGCCCCCTGG      900
E--> 1036 CCCCCGCCGG AAGGGGCCTT CGTGGGCTTT GTCCCTTTCCC GCCCCGAGCC
1037 CATGTGGGCC      960
E--> 1039 GAGCTTCTGG CCCTGGCCGC CGCCAGGGAG GGCCGGGTCC ACCGGGCACC
1040 AGACCCCTTT     1020
E--> 1042 ANGCGCCTNA GGGACCTNAA GGAGGTGCGG GGNCTCCTCG CCAAGGACCT
1043 GGCCGTTTTG     1080
E--> 1045 GCCCTGAGGG AGGGCCTNGA CCTCNTGCCC GGGGACGACC CCATGCTCCT
1046 CGCCTACCTC     1140
E--> 1048 CTGGACCCCT CCAACACCAC CCCCAGAGGG GTGGCCCGGC GCTACGGGGG
1049 GGAGTGGACG     1200
E--> 1051 GAGGANGCGG GGGAGCGGGC CCTCCTNTCC GAGAGGCTCT TCCNGAACCT
1052 NNNGCAGCGC     1260
E--> 1054 CTTGAGGGGG AGGAGAGGCT CTTTTGGCTT TACCAGGAGG TGGAGAAGCC
1055 CCTTTCCCGG     1320
E--> 1057 GTCCTGGCCC ACATGGAGGC CACGGGGGTN CGGCTGGACG TGGCCTACCT
1058 CCAGGCCCTN     1380
E--> 1060 TCCCTGGAGG TGGCGGAGGA GATCCGCCGC CTCGAGGAGG AGGTCTTCCG
1061 CCTGGCCGGC     1440
E--> 1063 CACCCCTTCA ACCTCAACTC CCGGGACCAG CTGGAAAGGG TGCTCTTTGA
1064 CGAGCTNNGG     1500
E--> 1066 CTTCCCGCCA TCGGCAAGAC GGAGAAGACN GGCAAGCGCT CCACCAGCGC
1067 CGCCGTGCTG     1560
E--> 1069 GAGGCCCTNC GNGAGGCCCA CCCCATCGTG GAGAAGATCC TGCAGTACCG
1070 GGAGCTCACC     1620
E--> 1072 AAGCTCAAGA ACACCTACAT NGACCCCTG CCNGNCCTCG TCCACCCAG
1073 GACGGGCCGC     1680
E--> 1075 CTCCACACCC GCTTCAACCA GACGGCCACG GCCACGGGCA GGCTTAGTAG
1076 CTCCGACCCC     1740
E--> 1078 AACCTGCAGA ACATCCCCGT CCGCACCCCN CTGGGCCAGA GGATCCGCCG
1079 GGCTTTCGTG     1800
E--> 1081 GCCGAGGAGG GNTGGGTGTT GGTGGCCCTG GACTATAGCC AGATAGAGCT
1082 CCGGGTCCTG     1860
E--> 1084 GCCCACCTCT CCGGGGACGA GAACCTGATC CGGGTCTTCC AGGAGGGGAG
1085 GGACATCCAC     1920
E--> 1087 ACCCAGACCG CCAGCTGGAT GTTCGGCGTC CCCCCGGAGG CCGTGGACCC
1088 CCTGATGCGC     1980
E--> 1090 CGGGCGGCCA AGACCATCAA CTTGCGGGTC CTCTACGGCA TGTCCGCCCA
1091 CCGCCTCTCC     2040

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 1093 CAGGAGCTTG CCATCCCCTA CGAGGAGGCG GTGGCCTTCA TTGAGCGCTA
      1094 CTTCCAGAGC      2100
E--> 1096 TTCCCCAAGG TGCGGGCCTG GATTGAGAAG ACCCTGGAGG AGGGCAGGAG
      1097 GCGGGGGTAC      2160
E--> 1099 GTGGAGACCC TCTTCGGCCG CCGGCGCTAC GTGCCCCGACC TCAACGCCCC
      1100 GGTGAAGAGC      2220
E--> 1102 GTGCGGGAGG CGGCGGAGCG CATGGCCTTC AACATGCCCC TCCAGGGCAC
      1103 CGCCGCCGAC      2280
E--> 1105 CTCATGAAGC TGGCCATGGT GAAGCTCTTC CCCC GGCTNC AGGAAATGGG
      1106 GGCCAGGATG      2340
E--> 1108 CTCCTNCAGG TCCACGACGA GCTGGTCCTC GAGGCCCCCA AAGAGCGGGC
      1109 GGAGGNGGTG      2400
E--> 1111 GCCGCTTTGG CCAAGGAGGT CATGGAGGGG GTCTATCCCC TGGCCGTGCC
      1112 CCTGGAGGTG      2460
E--> 1114 GAGGTGGGGA TGGGGGAGGA CTGGCTCTCC GCCAAGGAGT AG
W--> 1115 2502
      1290 (2) INFORMATION FOR SEQ ID NO: 9:
      1292 (i) SEQUENCE CHARACTERISTICS:
      1293 (A) LENGTH: 1647 base pairs
      1294 (B) TYPE: nucleic acid
      1295 (C) STRANDEDNESS: double
      1296 (D) TOPOLOGY: linear
      1298 (ii) MOLECULE TYPE: DNA (genomic)
      1302 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
E--> 1304 ATGAATTCGG GGATGCTGCC CCTCTTTGAG CCAAGGGGCC GGGTCCTCCT
      1305 GGTGGACGGC      60
E--> 1307 CACCACCTGG CCTACGCAC CTTCCACGCC CTGAAGGGCC TCACCACCAG
      1308 CCGGGGGGAG      120
E--> 1310 CCGGTGCAGG CGGTCTACGG CTTGCGCAAG AGCCTCCTCA AGGCCCTCAA
      1311 GGAGGACGGG      180
E--> 1313 GACGCGGTGA TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
      1314 GGCCTACGGG      240
E--> 1316 GGGTACAAGG CGGGCCGGGC CCCCACGCCG GAGGACTTTC CCCGGCAACT
      1317 CGCCCTCATC      300
E--> 1319 AAGGAGCTGG TGGACCTCCT GGGGCTGGCG CGCCTCGAGG TCCCGGGCTA
      1320 CGAGGCGGAC      360
E--> 1322 GACGTCTGG CCAGCCTGGC CAAGAAGGCG GAAAAGGAGG GCTACGAGGT
      1323 CCGCATCCTC      420
E--> 1325 ACCGCCGACA AAGACCTTTA CCAGCTCCTT TCCGACCGCA TCCACGTCCT
      1326 CCACCCCGAG      480
E--> 1328 GGGTACCTCA TCACCCGGC CTGGCTTTGG GAAAAGTACG GCCTGAGGCC
      1329 CGACCACTGG      540
E--> 1331 GCCGACTACC GGGCCCTGAC CGGGGACGAG TCCGACAACC TTCCCGGGGT
      1332 CAAGGGCATC      600
E--> 1334 GGGGAGAAGA CGGCGAGGAA GCTTCTGGAG GAGTGGGGGA GCCTGGAAGC
      1335 CCTCCTCAAG      660
E--> 1337 AACCTGGACC GGCTGAAGCC CGCCATCCGG GAGAAGATCC TGGCCCACAT
      1338 GGACGATCTG      720
E--> 1340 AAGCTCTCCT GGGACCTGGC CAAGGTGCGC ACCGACCTGC CCCTGGAGGT

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1341 GGACTTCGCC      780
E--> 1343 AAAAGGCGGG AGCCCGACCG GGAGAGGCTT AGGGCCTTTC TGGAGAGGCT
1344 TGAGTTTGGC      840
E--> 1346 AGCCTCCTCC ACGAGTTCGG CCTTCTGGAA AGCCCCAAGG CCCTGGAGGA
1347 GGCCCCCTGG      900
E--> 1349 CCCCCGCCGG AAGGGGCCTT CGTGGGCTTT GTGCTTTCCC GCAAGGAGCC
1350 CATGTGGGCC      960
E--> 1352 GATCTTCTGG CCCTGGCCGC CGCCAGGGGG GGCCGGGTCC ACCGGGCCCC
1353 CGAGCCTTAT     1020
E--> 1355 AAAGCCCTCA GGGACCTGAA GGAGGCGCGG GGGCTTCTCG CCAAAGACCT
1356 GAGCGTTCTG     1080
E--> 1358 GCCCTGAGGG AAGGCCTTGG CCTCCCGCCC GCGGACGACC CCATGCTCCT
1359 CGCCTACCTC     1140
E--> 1361 CTGGACCCTT CCAACACCAC CCCCAGAGGG GTGGCCCGGC GCTACGGCGG
1362 GGAGTGGACG     1200
E--> 1364 GAGGAGGCGG GGGAGCGGGC CGCCCTTTCC GAGAGGCTCT TCGCCAACCT
1365 GTGGGGGAGG     1260
E--> 1367 CTTGAGGGGG AGGAGAGGCT CCTTTGGCTT TACCGGGAGG TGGAGAGGCC
1368 CCTTTCCGCT     1320
E--> 1370 GTCCTGGCCC ACATGGAGGC CACGGGGGTG CGCCTGGACG TGGCCTATCT
1371 CAGGGCCTTG     1380
E--> 1373 TCCCTGGAGG TGGCCGGGGA GATCGCCCGC CTCGAGGCCG AGGTCTTCCG
1374 CCTGGCCGGC     1440
E--> 1376 CACCCCTTCA ACCTCAACTC CCGGGACCAG CTGGAAAGGG TCCTCTTTGA
1377 CGAGCTAGGG     1500
E--> 1379 CTTCCCGCCA TCGGCAAGAC GGAGAAGACC GGCAAGCGCT CCACCAGCGC
1380 CGCCGTCCTG     1560
E--> 1382 GAGGCCCTCC GCGAGGCCCA CCCCATCGTG GAGAAGATCC TGCAGGCATG
1383 CAAGCTTGGC     1620
1385 ACTGGCCGTC GTTTACAAC GTCGTGA
1387 (2) INFORMATION FOR SEQ ID NO: 10:
1389 (i) SEQUENCE CHARACTERISTICS:
1390 (A) LENGTH: 2088 base pairs
1391 (B) TYPE: nucleic acid
1392 (C) STRANDEDNESS: double
1393 (D) TOPOLOGY: linear
1395 (ii) MOLECULE TYPE: DNA (genomic)
1399 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
E--> 1401 ATGAATTCGG GGATGCTGCC CCTCTTTGAG CCAAGGGGCC GGGTCCTCCT
1402 GGTGGACGGC      60
E--> 1404 CACCACCTGG CCTACCGCAC CTTCCACGCC CTGAAGGGCC TCACCACCAG
1405 CCGGGGGGAG     120
E--> 1407 CCGGTGCAGG CGGTCTACGG CTTGCGCAAG AGCCTCCTCA AGGCCCTCAA
1408 GGAGGACGGG     180
E--> 1410 GACGCGGTGA TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
1411 GGCCTACGGG     240
E--> 1413 GGGTACAAGG CGGGCCGGGC CCCACGCCG GAGGACTTTC CCCGGCAACT
1414 CGCCCTCATC     300
E--> 1416 AAGGAGCTGG TGGACCTCCT GGGGCTGGCG CGCCTCGAGG TCCCGGGCTA

```

Same

1647

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1417 CGAGGCGGAC      360
E--> 1419 GACGTCCTGG CCAGCCTGGC CAAGAAGGCG GAAAAGGAGG GCTACGAGGT
1420 CCGCATCCTC      420
E--> 1422 ACCGCCGACA AAGACCTTTA CCAGCTCCTT TCCGACCGCA TCCACGTCCT
1423 CCACCCCGAG      480
E--> 1425 GGGTACCTCA TCACCCCGGC CTGGCTTTGG GAAAAGTACG GCCTGAGGCC
1426 CGACCACTGG      540
E--> 1428 GCCGACTACC GGGCCCTGAC CGGGGACGAG TCCGACAACC TTCCCGGGGT
1429 CAAGGGCATC      600
E--> 1431 GGGGAGAAGA CGGCGAGGAA GCTTCTGGAG GAGTGGGGGA GCCTGGAAGC
1432 CCTCCTCAAG      660
E--> 1434 AACCTGGACC GGCTGAAGCC CGCCATCCGG GAGAAGATCC TGGCCACAT
1435 GGACGATCTG      720
E--> 1437 AAGCTCTCCT GGGACCTGGC CAAGGTGCGC ACCGACCTGC CCCTGGAGGT
1438 GGA CTTCGCC      780
E--> 1440 AAAAGGCGGG AGCCCGACCG GGAGAGGCTT AGGGCCTTTC TGGAGAGGCT
1441 TGAGTTTGGC      840
E--> 1443 AGCCTCCTCC ACGAGTTTCG CCTTCTGGAA AGCCCCAAGG CCCTGGAGGA
1444 GGCCCCCTGG      900
E--> 1446 CCCCCGCCGG AAGGGGCCTT CGTGGGCTTT GTGCTTTCCC GCAAGGAGCC
1447 CATGTGGGCC      960
E--> 1449 GATCTTCTGG CCCTGGCCGC CGCCAGGGGG GGCCGGGTCC ACCGGGCCCC
1450 CGAGCCTTAT     1020
E--> 1452 AAAGCCCTCA GGGACCTGAA GGAGGCGCGG GGGCTTCTCG CCAAAGACCT
1453 GAGCGTTCTG     1080
E--> 1455 GCCCTGAGGG AAGGCCTTGG CCTCCCGCCC GCGGACGACC CCATGCTCCT
1456 CGCCTACCTC     1140
E--> 1458 CTGGACCCTT CCAACACCAC CCCCAGGGGG GTGGCCCGGC GCTACGGCGG
1459 GGAGTGGACG     1200
E--> 1461 GAGGAGGCGG GGGAGCGGGC CGCCCTTTCC GAGAGGCTCT TCGCCAACCT
1462 GTGGGGGAGG     1260
E--> 1464 CTTGAGGGGG AGGAGAGGCT CCTTTGGCTT TACCGGGAGG TGGAGAGGCC
1465 CCTTCCGCT      1320
E--> 1467 GTCTTGCCCC ACATGGAGGC CACGGGGGTG CGCCTGGACG TGGCCTATCT
1468 CAGGGCCTTG     1380
E--> 1470 TCCCTGGAGG TGGCCGGGGA GATCGCCCGC CTCGAGGCCG AGGTCTTCCG
1471 CCTGGCCGGC     1440
E--> 1473 CACCCCTTCA ACCTCAACTC CCGGGACCAG CTGGAAAGGG TCCTCTTTGA
1474 CGAGCTAGGG     1500
E--> 1476 CTTCCCGCCA TCGGCAAGAC GGAGAAGACC GGCAAGCGCT CCACCAGCGC
1477 CGCCGTCCTG     1560
E--> 1479 GAGGCCCTCC GCGAGGCCCA CCCCATCGTG GAGAAGATCC TGCAGTACCG
1480 GGAGCTCACC     1620
E--> 1482 AAGCTGAAGA GCACCTACAT TGACCCCTTG CCGGACCTCA TCCACCCAG
1483 GACGGGCCGC     1680
E--> 1485 CTCCACACCC GCTTCAACCA GACGGCCACG GCCACGGGCA GGCTAAGTAG
1486 CTCCGATCCC     1740
E--> 1488 AACCTCCAGA ACATCCCCGT CCGCACCCCG CTTGGGCAGA GGATCCGCCG
1489 GGCCTTCATC     1800

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 1491 GCCGAGGAGG GGTGGCTATT GGTGGCCCTG GACTATAGCC AGATAGAGCT
      1492 CAGGGTGCTG      1860
E--> 1494 GCCCACCTCT CCGGCGACGA GAACCTGATC CGGGTCTTCC AGGAGGGGCG
      1495 GGACATCCAC      1920
E--> 1497 ACGGAGACCG CCAGCTGGAT GTTCGGCGTC CCCCgggagg CCGTGGACCC
      1498 CCTGATGCGC      1980
E--> 1500 CGGGCGGCCA AGACCATCAA CTTCGGGGTC CTCTACGGCA TGTCGGCCCA
      1501 CCGCTCTCC      2040
E--> 1503 CAGGAGCTAG CTAGCCATCC CTTACGAGGA GGCCAGGCC TTCATTGA
W--> 1504 2088
      1506 (2) INFORMATION FOR SEQ ID NO: 11:
      1508 (i) SEQUENCE CHARACTERISTICS:
      1509 (A) LENGTH: 962 base pairs
      1510 (B) TYPE: nucleic acid
      1511 (C) STRANDEDNESS: double
      1512 (D) TOPOLOGY: linear
      1514 (ii) MOLECULE TYPE: DNA (genomic)
      1518 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:
E--> 1520 ATGAATTCGG GGATGCTGCC CCTCTTTGAG CCAAGGGCC GGGTCCTCCT
      1521 GGTGGACGGC      60
E--> 1523 CACCACCTGG CCTACCGCAC CTTCCACGCC CTGAAGGGCC TCACCACCAG
      1524 CCGGGGGGAG      120
E--> 1526 CCGGTGCAGG CGGTCTACGG CTTGCCAAG AGCCTCCTCA AGGCCCTCAA
      1527 GGAGGACGGG      180
E--> 1529 GACGCGGTGA TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
      1530 GGCTACGGG      240
E--> 1532 GGGTACAAGG CGGGCCGGGC CCCCACGCCG GAGGACTTTC CCCGGCAACT
      1533 CGCCCTCATC      300
E--> 1535 AAGGAGCTGG TGGACCTCCT GGGGCTGGCG CGCCTCGAGG TCCCGGGCTA
      1536 CGAGGCGGAC      360
E--> 1538 GACGTCTCTG CCAGCCTGGC CAAGAAGGCG GAAAAGGAGG GCTACGAGGT
      1539 CCGCATCCTC      420
E--> 1541 ACCGCCGACA AAGACCTTTA CCAGCTTCTT TCCGACCGCA TCCACGTCCT
      1542 CCACCCCGAG      480
E--> 1544 GGGTACCTCA TCACCCCGGC CTGGCTTTGG GAAAAGTACG GCCTGAGGCC
      1545 CGACCACTGG      540
E--> 1547 GCCGACTACC GGGCCCTGAC CGGGGACGAG TCCGACAACC TTCCCGGGGT
      1548 CAAGGGCATC      600
E--> 1550 GGGGAGAAGA CGGCGAGGAA GCTTCTGGAG GAGTGGGGGA GCCTGGAAGC
      1551 CCTCCTCAAG      660
E--> 1553 AACCTGGACC GGCTGAAGCC CGCCATCCGG GAGAAGATCC TGGCCACAT
      1554 GGACGATCTG      720
E--> 1556 AAGCTCTCCT GGGACCTGGC CAAGGTGCGC ACCGACCTGC CCCTGGAGGT
      1557 GGACTTCGCC      780
E--> 1559 AAAAGGCGGG AGCCCGACCG GGAGAGGCTT AGGGCCTTTC TGGAGAGGCT
      1560 TGAGTTTGGC      840
E--> 1562 AGCCTCCTCC ACGAGTTCCG CTTTCTGGAA AGCCCCAAGT CATGGAGGGG
      1563 GTGTATCCCC      900
E--> 1565 TGGCCGTGCC CCTGGAGGTG GAGGTGGGGA TAGGGGAGGA CTGGCTCTCC

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1566 GCCAAGGAGT      960
1568 GA
1570 (2) INFORMATION FOR SEQ ID NO: 12:
1572     (i) SEQUENCE CHARACTERISTICS:
1573         (A) LENGTH: 1600 base pairs
1574         (B) TYPE: nucleic acid
1575         (C) STRANDEDNESS: double
1576         (D) TOPOLOGY: linear
1578     (ii) MOLECULE TYPE: DNA (genomic)
1582     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:
E--> 1584 ATGGAATTCG GGGATGCTGC CCCTCTTTGA GCCCAAGGGC CGGGTCTCTCC
1585 TGGTGGACGG      60
E--> 1587 CCACCACCTG GCCTACCGCA CCTTCCACGC CCTGAAGGGC CTCACCACCA
1588 GCCGGGGGGA      120
E--> 1590 GCCGGTGCAG GCGGTCTACG GCTTCGCCAA GAGCCTCCTC AAGGCCCTCA
1591 AGGAGGACGG      180
E--> 1593 GGACGCGGTG ATCGTGGTCT TTGACGCCAA GGCCCCCTCC TTCCGCCACG
1594 AGGCCTACGG      240
E--> 1596 GGGGTACAAG GCGGGCCGGG CCCCCACGCC GGAGGACTTT CCCC GGCAAC
1597 TCGCCCTCAT      300
E--> 1599 CAAGGAGCTG GTGGACCTCC TGGGGCTGGC GCGCCTCGAG GTCCCGGGCT
1600 ACGAGGCGGA      360
E--> 1602 CGACGTCCTG GCCAGCCTGG CCAAGAAGGC GGAAAAGGAG GGCTACGAGG
1603 TCCGCATCCT      420
E--> 1605 CACCGCCGAC AAAGACCTTT ACCAGCTCCT TTCCGACCGC ATCCACGTCC
1606 TCCACCCCGA      480
E--> 1608 GGGGTACCTC ATCACCCCGG CCTGGCTTTG GGAAAAGTAC GGCCTGAGGC
1609 CCGACCAAGT      540
E--> 1611 GGCCGACTAC CGGGCCCTGA CCGGGGACGA GTCCGACAAC CTTCCCGGGG
1612 TCAAGGGCAT      600
E--> 1614 CGGGGAGAAG ACGGCGAGGA AGCTTCTGGA GGAGTGGGGG AGCCTGGAAG
1615 CCCTCCTCAA      660
E--> 1617 GAACCTGGAC CGGCTGAAGC CCGCCATCCG GGAGAAGATC CTGGCCCACA
1618 TGGACGATCT      720
E--> 1620 GAAGCTCTCC TGGGACCTGG CCAAGGTGCG CACCGACCTG CCCCTGGAGG
1621 TGGACTTCGC      780
E--> 1623 CAAAAGGCGG GAGCCCGACC GGGAGAGGCT TAGGGCCTTT CTGGAGAGGC
1624 TTGAGTTTGG      840
E--> 1626 CAGCCTCCTC CACGAGTTCG GCCTTCTGGA AAGCCCCAAG ATCCGCCGGG
1627 CCTTCATCGC      900
E--> 1629 CGAGGAGGGG TGGCTATTGG TGGCCCTGGA CTATAGCCAG ATAGAGCTCA
1630 GGGTGCTGGC      960
E--> 1632 CCACCTCTCC GGCACGAGA ACCTGATCCG GGTCTTCCAG GAGGGGCGGG
1633 ACATCCACAC      1020
E--> 1635 GGAGACCGCC AGCTGGATGT TCGGCGTCCC CCGGGAGGCC GTGGACCCCC
1636 TGATGCGCCG      1080
E--> 1638 GGCGGCCAAG ACCATCAACT TCGGGGTCCT CTACGGCATG TCGGCCACCC
1639 GCCTCTCCCA      1140
E--> 1641 GGAGCTAGCC ATCCCTTACG AGGAGGCCCA GGCCTTCATT GAGCGCTACT

```

962

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1642 TTCAGAGCTT      1200
E--> 1644 CCCCAAGGTG CGGGCCTGGA TTGAGAAGAC CCTGGAGGAG GGCAGGAGGC
1645 GGGGGTACGT      1260
E--> 1647 GGAGACCCTC TTCGGCCGCC GCCGCTACGT GCCAGACCTA GAGGCCCGGG
1648 TGAAGAGCGT      1320
E--> 1650 GCGGGAGGCG GCCGAGCGCA TGGCCTTCAA CATGCCCGTC CGGGGCACCG
1651 CCGCCGACCT      1380
E--> 1653 CATGAAGCTG GCTATGGTGA AGCTCTTCCC CAGGCTGGAG GAAATGGGGG
1654 CCAGGATGCT      1440
E--> 1656 CCTTCAGGTC CACGACGAGC TGGTCCTCGA GGCCCCAAAA GAGAGGGCGG
1657 AGGCCGTGGC      1500
E--> 1659 CCGGCTGGCC AAGGAGGTCA TGGAGGGGGT GATCCCCCTG GCCGTGCCCC
1660 TGGAGGTGGA      1560
E--> 1662 GGTGGGGATA GGGGAGGACT GGCTCTCCGC CAAGGAGTGA
W--> 1663 1600
1697 (2) INFORMATION FOR SEQ ID NO: 15:
1699 (i) SEQUENCE CHARACTERISTICS:
1700 (A) LENGTH: 91 base pairs
1701 (B) TYPE: nucleic acid
1702 (C) STRANDEDNESS: single
1703 (D) TOPOLOGY: linear
1705 (ii) MOLECULE TYPE: DNA (genomic)
1709 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:
E--> 1711 TAATACGACT CACTATAGGG AGACCGGAAT TCGAGCTCGC CCGGGCGAGC
1712 TCGAATTCCG      60
1714 TGTATTCTAT AGTGTCACCT AAATCGAATT C
1764 (2) INFORMATION FOR SEQ ID NO: 19:
1766 (i) SEQUENCE CHARACTERISTICS:
1767 (A) LENGTH: 42 base pairs
1768 (B) TYPE: nucleic acid
1769 (C) STRANDEDNESS: single
1770 (D) TOPOLOGY: linear
1772 (ii) MOLECULE TYPE: DNA (genomic)
1776 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:
E--> 1778 GGATCCTCTA GAGTCGACCT GCAGGCATGC CTACCTTGGT AG
W--> 1779 42
1797 (2) INFORMATION FOR SEQ ID NO: 21:
1799 (i) SEQUENCE CHARACTERISTICS:
1800 (A) LENGTH: 2502 base pairs
1801 (B) TYPE: nucleic acid
1802 (C) STRANDEDNESS: double
1803 (D) TOPOLOGY: linear
1805 (ii) MOLECULE TYPE: DNA (genomic)
1809 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:
E--> 1811 ATGAATTCGG GGATGCTGCC CCTCTTTGAG CCCAAGGGCC GGGTCCTCCT
1812 GGTGGACGGC      60
E--> 1814 CACCACCTGG CCTACCGCAC CTTCCACGCC CTGAAGGGCC TCACCACCAG
1815 CCGGGGGGAG      120
E--> 1817 CCGGTGCAGG CGGTCTACGG CTTGCCAAG AGCCTCCTCA AGGCCCTCAA

```

Same

91

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

1818 GGAGGACGGG      180
E--> 1820 GACGCGGTGA TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA
1821 GGCCTACGGG      240
E--> 1823 GGGTACAAGG CGGGCCGGGC CCCACGCCG GAGGACTTTC CCCGGCAACT
1824 CGCCCTCATC      300
E--> 1826 AAGGAGCTGG TGGACCTCCT GGGGCTGGCG CGCCTCGAGG TCCCGGGCTA
1827 CGAGGCGGAC      360
E--> 1829 GACGTCCTGG CCAGCCTGGC CAAGAAGGCG GAAAAGGAGG GCTACGAGGT
1830 CCGCATCCTC      420
E--> 1832 ACCGCCGACA AAGACCTTTA CCAGCTCCTT TCCGACCGCA TCCACGTCTT
1833 CCACCCCGAG      480
E--> 1835 GGGTACCTCA TCACCCCGGC CTGGCTTTGG GAAAAGTACG GCCTGAGGCC
1836 CGACCAAGTGG      540
E--> 1838 GCCGACTACC GGGCCCTGAC CGGGGACGAG TCCGACAACC TTCCCGGGGT
1839 CAAGGGCATC      600
E--> 1841 GGGGAGAAGA CGGCGAGGAA GCTTCTGGAG GAGTGGGGGA GCCTGGAAGC
1842 CCTCTCAAG      660
E--> 1844 AACCTGGACC GGCTGAAGCC CGCCATCCGG GAGAAGATCC TGGCCACAT
1845 GGACGATCTG      720
E--> 1847 AAGCTCTCCT GGGACCTGGC CAAGGTGCGC ACCGACCTGC CCCTGGAGGT
1848 GGACTTCGCC      780
E--> 1850 AAAAGGCGGG AGCCCGACCG GGAGAGGCTT AGGGCCTTTC TGGAGAGGCT
1851 TGAGTTTGGC      840
E--> 1853 AGCCTCCTCC ACGAGTTCGG CCTTCTGGAA AGCCCCAAGG CCCTGGAGGA
1854 GGCCCCCTGG      900
E--> 1856 CCCC GCCCGG AAGGGGCCCTT CGTGGGCTTT GTGCTTTCCC GCAAGGAGCC
1857 CATGTGGGCC      960
E--> 1859 GATCTTCTGG CCCTGGCCGC CGCCAGGGGG GGCCGGGTCC ACCGGGCCCC
1860 CGAGCCTTAT     1020
E--> 1862 AAAGCCCTCA GGGACCTGAA GGAGGCGCGG GGGCTTCTCG CCAAAGACCT
1863 GAGCGTCTG      1080
E--> 1865 GCCCTGAGGG AAGGCCTTGG CCTCCCGCCC GGCGACGACC CCATGCTCCT
1866 CGCTACCTC      1140
E--> 1868 CTGGACCCTT CCAACACCAC CCCCAGGGGG GTGGCCCGGC GCTACGGCGG
1869 GGAGTGGACG     1200
E--> 1871 GAGGAGGCGG GGGAGCGGGC CGCCCTTTCC GAGAGGCTCT TCGCCAACCT
1872 GTGGGGGAGG     1260
E--> 1874 CTTGAGGGGG AGGAGAGGCT CCTTTGGCTT TACCGGGAGG TGGAGAGGCC
1875 CCTTCCGCT      1320
E--> 1877 GTCCTGGCCC ACATGGAGGC CACGGGGGTG CGCCTGGACG TGGCCTATCT
1878 CAGGGCCTTG     1380
E--> 1880 TCCCTGGAGG TGGCCGGGGA GATCGCCCGC CTCGAGGCCG AGGTCTTCCG
1881 CCTGGCCGGC     1440
E--> 1883 CACCCCTTCA ACCTCAACTC CCGGGACCAG CTGGAAAGGG TCCTCTTTGA
1884 CGAGCTAGGG     1500
E--> 1886 CTTCCCGCCA TCGGCAAGAC GGAGAAGACC GGCAAGCGCT CCACCAGCGC
1887 CGCCGTCCTG     1560
E--> 1889 GAGGCCCTCC GCGAGGCCCA CCCCATCGTG GAGAAGATCC TGCAGTACCG
1890 GGAGCTCACC     1620

```

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

E--> 1892 AAGCTGAAGA GCACCTACAT TGACCCCTTG CCGGACCTCA TCCACCCCAG
 1893 GACGGGCCGC 1680
 E--> 1895 CTCCACACCC GCTTCAACCA GACGGCCACG GCCACGGGCA GGCTAAGTAG
 1896 CTCCGATCCC 1740
 E--> 1898 AACCTCCAGA ACATCCCCGT CCGCACCCCG CTTGGGCAGA GGATCCGCCG
 1899 GGCCTTCATC 1800
 E--> 1901 GCCGAGGAGG GGTGGCTATT GGTGGCCCTG GACTATAGCC AGATAGAGCT
 1902 CAGGGTGCTG 1860
 E--> 1904 GCCCACCTCT CCGGCGACGA GAACCTGATC CGGGTCTTCC AGGAGGGGCG
 1905 GGACATCCAC 1920
 E--> 1907 ACGGAGACCG CCAGCTGGAT GTTCGGCGTC CCCCGGGAGG CCGTGGACCC
 1908 CCTGATGCGC 1980
 E--> 1910 CGGGCGGCCA AGACCATCAA CTTCGGGGTC CTCTACGGCA TGTCGGCCCA
 1911 CCGCTCTCC 2040
 E--> 1913 CAGGAGCTAG CCATCCCTTA CGAGGAGGCC CAGGCCTTCA TTGAGCGCTA
 1914 CTTTCAGAGC 2100
 E--> 1916 TTCCCAAGG TCGGGCCCTG GATTGAGAAG ACCCTGGAGG AGGGCAGGAG
 1917 GCGGGGGTAC 2160
 E--> 1919 GTGGAGACCC TCTTCGGCCG CCGCCGCTAC GTGCCAGACC TAGAGGCCCG
 1920 GGTGAAGAGC 2220
 E--> 1922 GTGCGGGAGG CGGCCGAGCG CATGGCCTTC AACATGCCCC TCCGGGGCAC
 1923 CGCCGCCGAC 2280
 E--> 1925 CTCATGAAGC TGGCTATGGT GAAGCTCTTC CCCAGGCTGG AGGAAATGGG
 1926 GGCCAGGATG 2340
 E--> 1928 CTCCTTCAGG TCCACGACGA GCTGGTCCTC GAGGCCCCAA AAGAGAGGGC
 1929 GGAGGCCGTG 2400
 E--> 1931 GCCCGGCTGG CCAAGGAGGT CATGGAGGGG GTGTATCCCC TGGCCGTGCC
 1932 CCTGGAGGTG 2460
 E--> 1934 GAGGTGGGGA TAGGGGAGGA CTGGCTCTCC GCCAAGGAGT GA
 W--> 1935 2502

Same

1953 (2) INFORMATION FOR SEQ ID NO: 23:
 1955 (i) SEQUENCE CHARACTERISTICS:
 1956 (A) LENGTH: 72 base pairs
 1957 (B) TYPE: nucleic acid
 1958 (C) STRANDEDNESS: single
 1959 (D) TOPOLOGY: linear
 1961 (ii) MOLECULE TYPE: DNA (genomic)
 1965 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:
 E--> 1967 CGGACGAACA AGCGAGACAG CGACACAGGT ACCACATGGT ACAAGAGGCA
 1968 AGAGAGACGA 60
 1970 CACAGCAGAA AC
 1972 (2) INFORMATION FOR SEQ ID NO: 24:
 1974 (i) SEQUENCE CHARACTERISTICS:
 1975 (A) LENGTH: 70 base pairs
 1976 (B) TYPE: nucleic acid
 1977 (C) STRANDEDNESS: single
 1978 (D) TOPOLOGY: linear
 1980 (ii) MOLECULE TYPE: DNA (genomic)
 1984 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

72

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

E--> 1986 GTTCTGCTG TGTCGTCTCT CTTGCCTCTT GTACCATGTG GTACCTGTGT
 1987 CGCTGTCTCG 60
 1989 CTTGTTCGTC
 2023 (2) INFORMATION FOR SEQ ID NO: 27:
 2025 (i) SEQUENCE CHARACTERISTICS:
 2026 (A) LENGTH: 46 base pairs
 2027 (B) TYPE: nucleic acid
 2028 (C) STRANDEDNESS: single
 2029 (D) TOPOLOGY: linear
 2031 (ii) MOLECULE TYPE: DNA (genomic)
 2035 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:
 E--> 2037 CCTCTGTAC CATGTGGTAC CTGTGTCGCT GTCTCGCTTG TTCGTC
 W--> 2038 46
 2040 (2) INFORMATION FOR SEQ ID NO: 28:
 2042 (i) SEQUENCE CHARACTERISTICS:
 2043 (A) LENGTH: 50 base pairs
 2044 (B) TYPE: nucleic acid
 2045 (C) STRANDEDNESS: single
 2046 (D) TOPOLOGY: linear
 2048 (ii) MOLECULE TYPE: DNA (genomic)
 2052 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:
 E--> 2054 ACACAGGTAC CACATGGTAC AAGAGGCAAG AGAGACGACA CAGCAGAAAC
 W--> 2055 50
 2074 (2) INFORMATION FOR SEQ ID NO: 30:
 2076 (i) SEQUENCE CHARACTERISTICS:
 2077 (A) LENGTH: 969 base pairs
 2078 (B) TYPE: nucleic acid
 2079 (C) STRANDEDNESS: double
 2080 (D) TOPOLOGY: linear
 2082 (ii) MOLECULE TYPE: DNA (genomic)
 2086 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:
 E--> 2088 ATGGCTAGCA TGACTGGTGG ACAGCAAATG GGTCCGATCA ATTCGGGGAT
 2089 GCTGCCCTC 60
 E--> 2091 TTGAGCCCC AGGGCCGGGT CCTCCTGGTG GACGGCCACC ACCTGGCCTA
 2092 CCGCACCTTC 120
 E--> 2094 CACGCCCTGA AGGGCCTCAC CACCAGCCGG GGGGAGCCGG TGCAGGCGGT
 2095 CTACGGCTTC 180
 E--> 2097 GCCAAGAGCC TCCTCAAGGC CCTCAAGGAG GACGGGGACG CGGTGATCGT
 2098 GGTCTTTGAC 240
 E--> 2100 GCCAAGGCCC CCTCCTTCCG CCACGAGGCC TACGGGGGGT ACAAGGCGGG
 2101 CCGGGCCCCC 300
 E--> 2103 ACGCCGGAGG ACTTTCCCCG GCAACTCGCC CTCATCAAGG AGCTGGTGGG
 2104 CCTCCTGGGG 360
 E--> 2106 CTGGCGCGCC TCGAGGTCCC GGGCTACGAG GCGGACGACG TCCTGGCCAG
 2107 CCTGGCCAAG 420
 E--> 2109 AAGGCGGAAA AGGAGGGCTA CGAGGTCCGC ATCCTCACCG CCGACAAAGA
 2110 CCTTTACCAG 480
 E--> 2112 CTTCTTTCCG ACCGCATCCA CGTCCTCCAC CCCGAGGGGT ACCTCATCAC
 2113 CCCGGCCTGG 540

70

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 2115 CTTTGGGAAA AGTACGGCCT GAGGCCCCGAC CAGTGGGCCG ACTACCGGGC
      2116 CTGACCGGG      600
E--> 2118 GACGAGTCCG ACAACCTTCC CGGGGTCAAG GGCATCGGGG AGAAGACGGC
      2119 GAGGAAGCTT      660
E--> 2121 CTGGAGGAGT GGGGGAGCCT GGAAGCCCTC CTCAAGAACC TGGACCGGCT
      2122 GAAGCCCGCC      720
E--> 2124 ATCCGGGAGA AGATCCTGGC CCACATGGAC GATCTGAAGC TCTCCTGGGA
      2125 CCTGGCCAAG      780
E--> 2127 GTGCGCACCG ACCTGCCCCCT GGAGGTGGAC TTCGCCAAAA GCGGGGAGCC
      2128 CGACCGGGAG      840
E--> 2130 AGGCTTAGGG CCTTCTTGGA GAGGCTTGAG TTTGGCAGCC TCCTCCACGA
      2131 GTTCGGCCTT      900
E--> 2133 CTGGAAAGCC CCAAGTCATG GAGGGGGTGT ATCCCCTGGC CGTGCCCCCTG
      2134 GAGGTGGAGG      960
      2136 TGGGGATAG
      2138 (2) INFORMATION FOR SEQ ID NO: 31:
      2140 (i) SEQUENCE CHARACTERISTICS:
      2141 (A) LENGTH: 948 base pairs
      2142 (B) TYPE: nucleic acid
      2143 (C) STRANDEDNESS: double
      2144 (D) TOPOLOGY: linear
      2146 (ii) MOLECULE TYPE: DNA (genomic)
      2150 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:
E--> 2152 ATGGCTAGCA TGACTGGTGG ACAGCAAATG GGTCCGATCA ATTCGGGGAT
      2153 GCTGCCCCCTC      60
E--> 2155 TTTGAGCCCA AGGGCCGGGT CCTCCTGGTG GACGGCCACC ACCTGGCCTA
      2156 CCGCACCTTC      120
E--> 2158 CACGCCCTGA AGGGCCTCAC CACCAGCCGG GGGGAGCCGG TGCAGGCGGT
      2159 CTACGGCTTC      180
E--> 2161 GCCAAGAGCC TCCTCAAGGC CCTCAAGGAG GACGGGGACG CGGTGATCGT
      2162 GGTCTTTGAC      240
E--> 2164 GCCAAGGCC CCTCCTTCCG CCACGAGGCC TACGGGGGGT ACAAGGCGGG
      2165 CCGGGCCCCC      300
E--> 2167 ACGCCGGAGG ACTTTCCCG GCAACTCGCC CTCATCAAGG AGCTGGTGGA
      2168 CCTCCTGGGG      360
E--> 2170 CTGGCGCGCC TCGAGGTCCC GGGCTACGAG GCGGACGACG TCCTGGCCAG
      2171 CCTGGCCAAG      420
E--> 2173 AAGGCGGAAA AGGAGGGCTA CGAGGTCCGC ATCCTCACC GCGACAAAGA
      2174 CCTTTACCAG      480
E--> 2176 CTTCTTTCCG ACCGCATCCA CGTCCTCCAC CCCGAGGGGT ACCTCATCAC
      2177 CCCGGCCTGG      540
E--> 2179 CTTTGGGAAA AGTACGGCCT GAGGCCCCGAC CAGTGGGCCG ACTACCGGGC
      2180 CTGACCGGG      600
E--> 2182 GACGAGTCCG ACAACCTTCC CGGGGTCAAG GGCATCGGGG AGAAGACGGC
      2183 GAGGAAGCTT      660
E--> 2185 CTGGAGGAGT GGGGGAGCCT GGAAGCCCTC CTCAAGAACC TGGACCGGCT
      2186 GAAGCCCGCC      720
E--> 2188 ATCCGGGAGA AGATCCTGGC CCACATGGAC GATCTGAAGC TCTCCTGGGA
      2189 CCTGGCCAAG      780

```

969

Same

RAW SEQUENCE LISTING

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:44

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

```

E--> 2191 GTGCGCACCG ACCTGCCCCT GGAGGTGGAC TTCGCCAAAA GCGGGGAGCC
      2192 CGACCGGGAG      840
E--> 2194 AGGCTTAGGG CCTTTCTGGA GAGGCTTGAG TTTGGCAGCC TCCTCCACGA
      2195 GTTCGGCCTT      900
E--> 2197 CTGGAAAGCC CCAAGGCCGC ACTCGAGCAC CACCACCACC ACCACTGA
W--> 2198 948
      2200 (2) INFORMATION FOR SEQ ID NO: 32:
      2202 (i) SEQUENCE CHARACTERISTICS:
      2203 (A) LENGTH: 206 base pairs
      2204 (B) TYPE: nucleic acid
      2205 (C) STRANDEDNESS: double
      2206 (D) TOPOLOGY: linear
      2208 (ii) MOLECULE TYPE: DNA (genomic)
      2212 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:
E--> 2214 CGCCAGGGTT TTCCAGTCA CGACGTTGTA AAACGACGGC CAGTGAATTG
      2215 TAATACGACT      60
E--> 2217 CACTATAGGG CGAATTCGAG CTCGGTACCC GGGGATCCTC TAGAGTCGAC
      2218 CTGCAGGCAT      120
E--> 2220 GCAAGCTTGA GTATTCTATA GTGTCACCTA AATAGCTTGG CGTAATCATG
      2221 GTCATAGCTG      180
      2223 TTTCCTGTGT GAAATTGTTA TCCGCT
      2289 (2) INFORMATION FOR SEQ ID NO: 37:
      2291 (i) SEQUENCE CHARACTERISTICS:
      2292 (A) LENGTH: 43 base pairs
      2293 (B) TYPE: nucleic acid
      2294 (C) STRANDEDNESS: single
      2295 (D) TOPOLOGY: linear
      2297 (ii) MOLECULE TYPE: DNA (genomic)
      2301 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:
E--> 2303 GACGAACAAG CGAGACAGCG ACCAGAGAGC AGAGAACCCA GAA
W--> 2304 43
      2338 (2) INFORMATION FOR SEQ ID NO: 40:
      2340 (i) SEQUENCE CHARACTERISTICS:
      2341 (A) LENGTH: 157 base pairs
      2342 (B) TYPE: nucleic acid
      2343 (C) STRANDEDNESS: double
      2344 (D) TOPOLOGY: linear
      2346 (ii) MOLECULE TYPE: DNA (genomic)
      2350 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:
E--> 2352 CACCGTCCTC TTCAAGAAGT TTATCCAGAA GCCAATGCAC CCATTGGACA
      2353 TAACCGGGAA      60
E--> 2355 TCCTACATGG TTCCTTTTAT ACCACTGTAC AGAAATGGTG ATTTCTTTAT
      2356 TTCATCCAAA      120
      2358 GATCTGGGCT ATGACTATAG CTATCTACAA GATTCAG

```

Same

206

157



Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/940,925

DATE: 03/15/2002

TIME: 14:41:46

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:58 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:1
M:254 Repeated in SeqNo=1
L:182 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:1
L:198 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:2
M:254 Repeated in SeqNo=2
L:337 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:3
M:254 Repeated in SeqNo=3
L:461 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:991 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:7
M:254 Repeated in SeqNo=7
L:1115 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:7
L:1131 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1149 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1164 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1170 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1173 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1176 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1179 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1185 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1191 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1194 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1197 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1206 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1209 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1233 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1242 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1275 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1278 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1284 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1287 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8
L:1304 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:9
M:254 Repeated in SeqNo=9
L:1401 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:10
M:254 Repeated in SeqNo=10
L:1504 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:10
L:1520 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:11
M:254 Repeated in SeqNo=11
L:1584 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:12
M:254 Repeated in SeqNo=12
L:1663 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:12
L:1711 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:15
L:1778 M:254 E: No. of Bases conflict, Input:0 Counted:42 SEQ:19
L:1779 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:19
L:1811 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:21

VERIFICATION SUMMARY

DATE: 03/15/2002

PATENT APPLICATION: US/09/940,925

TIME: 14:41:46

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

M:254 Repeated in SeqNo=21

L:1935 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:21

L:1967 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:23

L:1986 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:24

L:2037 M:254 E: No. of Bases conflict, Input:0 Counted:46 SEQ:27

L:2038 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:27

L:2054 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:28

L:2055 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:28

L:2088 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:30

M:254 Repeated in SeqNo=30

L:2152 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:31

M:254 Repeated in SeqNo=31

L:2198 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:31

L:2214 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:32

M:254 Repeated in SeqNo=32

L:2303 M:254 E: No. of Bases conflict, Input:0 Counted:43 SEQ:37

L:2304 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:37

L:2352 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:40

M:254 Repeated in SeqNo=40

L:2374 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:41

M:254 Repeated in SeqNo=41

L:2428 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:44

M:254 Repeated in SeqNo=44

L:2491 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:47

M:254 Repeated in SeqNo=47

L:2513 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:48

M:254 Repeated in SeqNo=48

L:2535 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:49

M:254 Repeated in SeqNo=49

L:2542 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:49

L:2558 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:50

M:254 Repeated in SeqNo=50

L:2631 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:54

M:254 Repeated in SeqNo=54

L:2653 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:55

M:254 Repeated in SeqNo=55

L:2675 M:254 E: No. of Bases conflict, Input:0 Counted:50 SEQ:56

L:3255 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:70

L:3421 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:76

L:3515 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:79

L:3547 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:80

L:3579 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:81

L:4534 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:92

L:4608 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:93

L:4682 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:94

L:5487 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:121

L:5550 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:123

L:5579 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:124

L:5608 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:125

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/940,925

DATE: 03/15/2002

TIME: 14:41:46

Input Set : A:\Seq-sub.app

Output Set: N:\CRF3\03152002\I940925.raw

L:5637 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:126
L:5666 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:127
L:5695 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:128
L:5724 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:129
L:5753 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:130
L:5782 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:131
L:5811 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:132
L:6723 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:158
L:6924 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:161
L:7017 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:165